172, 400

The Catholic University of America
Washington, D. C.

Department of Mathematics
STATISTICAL LABORATORY

C.U.-Biom.-5
June 1963

AD No. 409104
DDC FILE COPY

RUNS TEST FOR A CIRCULAR DISTRIBUTION
AND A TABLE OF PROBABILITIES

by Chooichiro Asano

409104

JUL 1

~~Prepared Under Contract~~
~~No.~~ Nonr-2249(05)
U.S. Office of Naval Research

$1.60

RUNS TEST FOR A CIRCULAR DISTRIBUTION AND A TABLE OF PROBABILITIES.*

by

Chooichiro Asano **

- - -

0. Summary.

A method is suggested for testing whether two samples observed on a circle are drawn from the same distribution. The proposed test is a modification of the well-known Wald-Wolfowitz runs test for a distribution on a straight line. The primary advantage of the proposed test is that it minimizes the number of assumptions on the theoretical distribution.

1. Introduction.

Circular statistical problems arise in many scientific fields such as research on orientation of animals (see for instance Schmidt-Koenig (1960)), time period analysis for biological clocks, and rock magnetism in geology.

As Curray (1956) pointed out, for very large samples the two-sample problem can be treated by the use of the χ^2-test There is, however, a great need for a test that can be used when, due to the small size of the samples, the χ^2-test is not applicable.

In 1956 Watson and Williams proposed several two-sample tests based on the von Mises distribution (the so-called circular normal distribution). These tests, however, are all parametric. In applications there is not always evidence of the circular normal distribution.

---

* This study was supported in part by the contract NSF-9968, National Science Foundation; in part by the contract Nonr 2249(05),(Nr 301-579), the Office of Naval Research, Dept. of the Navy, with The Catholic University of America.

** This paper was written while the author was a Research Associate in the Statistical Laboratory, Department of Mathematics, The Catholic University of America, during the academic year 1962.

There exist also some non-parametric two-sample tests. One of them, proposed by Kuiper (1960), is a modification of the Kolmogorov-Smirnov test for a circle. Another such test is due to Watson (1962) who applies a similar procedure to a test suggested by Smirnov. At present the usefulness of these tests is limited since tables for the significant values of the test statistic are only partly available. Finally, a modification of the Mann-Whitney-Cochran test for a circle has been studied by Batschelet. The result is still unpublished.

The purpose of this paper is to propose a runs test for a circular distribution in the two-sample case which will be applicable when the sample size is small and which will involve a minimum number of assumptions on the theoretical distribution. This at first appeared to be only a matter of extending the considerations of the ordinary theory of runs on a line as proposed by Stevens (1939), Wald and Wolfowitz (1940), Mood (1940) Swed and Eisenhart (1943), and Wolfowitz (1943). But it soon became evident that further investigations were needed. The rotatable symmetry for the circle changes the mathematical treatment essentially. Where two separate sets of observations are combined in all the various possible ways and rotatable symmetry is considered, the resulting arrangements were sometimes identical with other cases already obtained. Hence, we must reduce the number of cases by the number of arrangements found to be identical with one already obtained.

We suppose that the merit and the properties of this test are quite similar to those of the ordinary runs test studied previously for points on a line. For practical applications the runs test is extremely simple and fast. The theoretical treatment of the runs test has the advantage that no discussion is necessary to justify the independence of the starting point. Research on the power of the circular runs test has not yet begun.

The numerical table of the distribution function of the test statistic was computed on an IBM 1620.

2. Probability function for the number of runs on a circle.

Given two sets of samples on a circle, we wish to determine the probability function for the number of runs obtainable by combining these two sets of samples. As a preliminary step we consider the following partition problem.

Suppose that there are k intervals on a circle and that in each interval there are n_i elements of the first sample, where the n_i's are ordered as follows:

$$(2.1) \qquad n_1 \geq n_2 \geq \cdots\cdots \geq n_k > 0 , \qquad \text{for } 1 \leq k \leq N ,$$

where $\sum_{i=1}^{k} n_i = N$ and N is now fixed. After the consideration of all possible arrangements of such k partitioned integers on a circle, we will consider how to combine the second sample among these k partitioned integers of the first sample.

Let

$$(2.2) \qquad S_{(k)} = \{n_1, n_2, \ldots\ldots, n_k \mid n_i \geq n_{i+1}, \sum_i n_i = N\}$$

denote the set or sets of ordered integers determined by the sample size N and the partition size k ; e.g., for $N = 4$ we have $S_{(1)} = \{4\}$, $S_{(2)} = \{3,1\}$, $\{2,2\}$, $S_{(3)} = \{2,1,1\}$, $S_{(4)} = \{1,1,1,1\}$. Thus the set $\{S_{(k)}\}$ $k=1,2,\ldots,N$ contains all the possible partitions for the integer N .

We now proceed to determine the number of rotatable symmetries generated from arranging these k partitioned integers on a circle.

If we introduce g_j such that

$$(2.3) \qquad g_j \equiv (\text{the number of } i\text{'s} \mid n_i = q,\ i = 1,2,\ldots,k) \quad \text{for } j=1,2,\ldots,t\ ,$$

where $n_k \leq q \leq n_1 \leq N$, $1 \leq t \leq k$, we can characterize each element in the set $\{S_{(k)}\}$ $k = 1,2,\ldots,N$. Then, corresponding to each element in $S_{(k)}$, we obtain a new set $G_{(k)}$ given by

$$(2.4) \qquad G_{(k)} = \{g_1,g_2,\ldots\ldots,g_t\} \quad .$$

Furthermore, let the common divisors, including one, of such $g_1,g_2,\ldots,g_t$ be $d_1,d_2,\ldots,d_p$, where we put $d_1 = 1<d_2<\ldots\ldots<d_p$ and $1 \leq p \leq t$.

If we omit repeated identical arrangement, then the number of circular permutations of the k integers is given by

$$(2.5) \qquad \Phi(n_1,n_2,\ldots,n_k|N) = \frac{1}{k}\sum_{i=1}^{p} \phi(d_i)\left(\frac{k}{d_i}\right)! \Big/ \left(\frac{g_1}{d_i}\right)!\left(\frac{g_2}{d_i}\right)!\ldots\left(\frac{g_p}{d_i}\right)!$$

where Euler's totient function $\phi(d_i)$, the number of positive integers less than d_i and prime to d_i , is given by

$$\phi(d_i) = d_i(1 - \frac{1}{p_1} \times 1 - \frac{1}{p_2}) \ \ldots .$$

and $p_1,p_2,\ldots.$ are the different prime factors of d_i , with the exception that $\phi(0) \equiv \phi(1) \equiv 1$. This corresponds to Barton and David (1958, 1962).

Similarly, the number $\Psi(n_1,n_2,\ldots,n_k|d_i,N)$ of symmetrical cases generated by $(\frac{d_i}{k})$-rotation# is obtained in the following way.

#The notation $(\frac{d_i}{k})$-rotation means a rotation of $(360 \times \frac{d_i}{k})^{o}$.

Let $m_1, m_2, \ldots, m_t$ denote the integers $\frac{g_1}{d_i}, \frac{g_2}{d_i}, \ldots, \frac{g_t}{d_i}$, respectively, and let $\Phi(m_1, m_2, \ldots, m_t | d_i, N)$ be given by

(2.7)
$$\Phi(m_1, m_2, \ldots, m_t | d_i, N) = \frac{1}{t} \sum_{d^*} \phi(d^*) \left(\frac{t}{d^*}\right)! \Big/ \left(\frac{m_1}{d^*}\right)! \left(\frac{m_2}{d^*}\right)! \cdots \left(\frac{m_t}{d^*}\right)! \quad ,$$

where the d^*'s denote the common divisors of $m_1, m_2, \ldots, m_t$, including one, and $\phi(d^*)$ is Euler's totient function, with the exception that $\phi(0) \equiv \phi(1) \equiv 1$.

Then the value of $\Psi(n_1, n_2, \ldots, n_k \mid d_i, N)$ is given by

(2.8)
$$\Psi(n_1, n_2, \ldots, n_k | d_i, N) = \Phi(m_1, m_2, \ldots, m_t | d_i, N) - \sum_{p \geq j \geq i+1}^{*} \Psi(n_1, n_2, \ldots, n_k | d_j, N) \; ,$$

where Σ^* means a summation over all multiples of d_i among $d_1, d_2, \ldots, d_p$. Naturally, as a result of (2.3), we obtain

(2.9)
$$\Psi(n_1, n_2, \ldots, n_k | d_i = 1, N) = 0 \; .$$

Now, on the basis of the above results, let us consider the probability function for the number of runs obtained by arranging simultaneously two different sets of observations on a circle.

Let N_1 and N_2 be the sizes of the first and second sample, respectively, where we assume $N_1 \leq N_2$ without any loss of generality.

First, the number $I(N_1, N_2)$ of all possible runs observable on a circle is obtained by using Euler's totient function again as follows:

(2.10)
$$I(N_1, N_2) = \frac{1}{N_1 + N_2} \sum_{d} \phi(d) \left(\frac{N_1 + N_2}{d}\right)! \Big/ \left(\frac{N_1}{d}\right)! \left(\frac{N_2}{d}\right)! \quad ,$$

where the summation is over all divisors, including one, of the greatest common divisor of N_1 and N_2 . This number will be the denominator in the determination of the probabilities.

Second, in order to give the probability function for an arbitrary

"2k runs" , $1 \leq k \leq N_1$, let us determine the number of possibilities that 2k runs will be observed. Let $n_1^{(s)}, n_2^{(s)}, \ldots, n_k^{(s)}$ be partitioned integers for N_s , $s=1,2$. Then the enumerators of both non-symmetrical arrangements for N_1 and N_2 on a circle are given by

$$(2.11) \quad \Phi(n_1^{(s)}, n_2^{(s)}, \ldots, n_k^{(s)} | N_s) - \sum_{i=1}^{p} \Psi(n_1^{(s)}, n_2^{(s)}, \ldots, n_k^{(s)} | d_i^{(s)}, N)$$

for $s = 1,2$, respectively, where $d_i^{(s)}$ indicates the d_i defined previously for s-th sample.

Hence, the number of arrangements obtainable by combining both non-symmetrical circular permutations for N_1 and N_2 is

(2.12)

$$I_1 \equiv k \prod_{s=1}^{2} \{\Phi(n_1^{(s)}, n_2^{(s)}, \ldots, n_k^{(s)} | N_s) - \sum_{i=1}^{p} \Psi(n_1^{(s)}, n_2^{(s)}, \ldots, n_k^{(s)} | d_i^{(s)}, N)\} .$$

Furthermore, where the symmetrical circular permutations of k integers for N_1 are combined with the non-symmetrical circular permutations for N_2 , the number of possible arrangements is given by

$$(2.13) \qquad I_2 \equiv \{ \sum_{d_i^{(1)}} \left(\frac{k}{d_i^{(1)}}\right) \Psi(n_1^{(1)}, n_2^{(1)}, \ldots, n_k^{(1)} | d_i^{(1)}, N_1)\}$$

$$\cdot\{\Phi({}_1^{(2)}, n_2^{(2)}, \ldots, n_k^{(2)} | N_2) - \sum_{i=1}^{p} \Psi(n_1^{(2)}, n_2^{(2)}, \ldots, n_k^{(2)} | d_i^{(2)}, N_2)\} .$$

Similarly, reversing the roles of N_1 and N_2 , we obtain

$$(2.14) \qquad I_3 \equiv \sum_{d_i^{(2)}} \left(\frac{k}{d_i^{(2)}}\right) \Psi(n_1^{(2)}, n_2^{(2)}, \ldots, n_k^{(2)} | d_i^{(2)}, N_2)\}$$

$$\cdot\{\Phi(n_1^{(1)}, n_2^{(1)}, \ldots, n_k^{(1)} | N_1) - \sum_{i=1}^{p} \Psi(n_1^{(1)}, n_2^{(1)}, \ldots, n_k^{(1)} | d_i^{(1)}, N_1)\} .$$

Finally, when both symmetrical circular permutations for N_1 and N_2 are combined, the number of possible arrangements is given by

$$(2.15) \quad I_4 \equiv \sum_{d_i^{(1)}} \sum_{d_j^{(2)}} G.C.D.\left(\frac{k}{d_i^{(1)}}, \frac{k}{d_j^{(2)}}\right) \Psi(n_1^{(1)}, n_2^{(1)}, \ldots, n_k^{(1)} | d_i^{(1)}, N_1) \cdot \Psi(n_1^{(2)}, n_2^{(2)}, \ldots, n_k^{(2)} | d_j^{(2)}, N_2) ,$$

where G.C.D. $\left(\frac{k}{d_i^{(1)}}, \frac{k}{d_j^{(2)}}\right)$ indicates the grestest common divisor of $\frac{k}{d_i^{(1)}}$ and $\frac{k}{d_j^{(2)}}$.

Thus putting

$$(2.16) \qquad I(N_1, N_2, k) = \sum_{\{S_{(k)}^{(1)}\}} \sum_{\{S_{(k)}^{(2)}\}} \sum_{u=1}^{4} I_u$$

we obtain the probability function of observing $2k$ runs by arranging simultaneously N_1 and N_2 observations on a circle, where $\{S_{(k)}^{(s)}\}$ indicates the set $S_{(k)}$ defined by (2.2) for $s=1,2$. This function is given by

$$(2.17) \quad P_r(k | N_1, N_2) = I(N_1, N_2, k) / I(N_1, N_2) \qquad \text{for } 1 \leq k \leq N_1 .$$

3. Distribution function and the table.

Now we can easily obtain the distribution function for the number of runs on a circle by using (2.17).

If $2k$ is defined to be the number of runs, then the probability of an arrangement yielding $2r$ or fewer runs is

$$(3.1) \qquad P_r\{2k \leq 2r\} = \sum_{k=1}^{r} P_r\{k | N_1, N_2\} .$$

The following table has been prepared for use in testing whether or not two sets of observations are from the same population. The Table

gives $P_r\{2k \leq 2r\}$ to 5 decimal places for $N_1 \leq N_2 \leq 20$ with a range of N_1 from 2 to 20.

TABLE of $P\{2r \leq 2k\}$, where $2r$ is the number of runs.

N_2	k	$N_1=2$	$N_1=3$	$N_1=4$	$N_1=5$	$N_1=6$	$N_1=7$	$N_1=8$	$N_1=9$
2	1	.5							
	2	1.							
3	1	.5	.25						
	2	1.	.75						
	3		1.						
4	1	.33333	.2	.1					
	2	1.	.8	.6					
	3		1.	.9					
	4			1.					
5	1	.33333	.14286	.07143	.03846				
	2	1.	.71429	.50000	.34615				
	3		1.	.92857	.80769				
	4			1.	.96154				
	5				1.				
6	1	.25	.1	.04545	.02381	.01250			
	2	1.	.6	.40909	.26190	.17500			
	3		1.	.86364	.73810	.60000			
	4			1.	.97619	.92500			
	5				1	.98750			
	6					1.			
7	1	.25	.08333	.03333	.01515	.00758	.00407		
	2	1.	.58333	.33333	.19697	.06897	.07724		
	3		1.	.83333	.65152	.50000	.38211		
	4			1.	.95455	.87879	.78862		
	5				1.	.99242	.97154		
	6					1.	.99593		
	7						1.		
8	1	.2	.06666	.02326	.01010	.00461	.00233	.00124	
	2	1.	.53333	.27907	.15152	.08756	.05128	.03218	
	3		1.	.76744	.57576	.41014	.29604	.21411	
	4			1.	.92930	.82028	.70396	.59406	
	5				1.	.98157	.94872	.89728	
	6					1.	.99767	.99010	
	7						1.	.99876	
	8							1.	
9	1	.2	.05263	.01818	.00699	.00299	.00140	.00070	.00037
	2	1.	.47368	.23636	.11889	.06269	.03497	.02028	.01224
	3		1.	.74545	.51049	.34328	.23077	.15734	.10612
	4			1.	.90210	.76119	.62238	.50000	.41233
	5				1.	.97015	.91608	.84266	.76067
	6					1.	.99441	.97972	.95510
	7						1.	.99930	.99667
	8							1.	.99963
	9								1.

Table of P{2r ≤ 2k} (continued).

N_2	k	N_1=2	N_1=3	N_1=4	N_1=5	N_1=6	N_1=7	N_1=8	N_1=9
10	1	.16667	.04545	.01370	.00498	.00198	.00086	.00041	.00021
	2	1.	.45455	.20548	.09453	.04762	.02405	.01354	.00761
	3		1.	.69863	.45274	.28571	.17869	.11738	.07672
	4			1.	.87065	.70635	.53952	.42133	.31859
	5				1.	.95635	.86426	.78460	.68141
	6					1.	.98969	.96344	.92328
	7						1.	.99794	.99238
	8							1.	.99979
	9								1.
	10								
11	1	.16667	.03846	.01099	.00366	.00137	.00057	.00025	.00012
	2	1.	.42308	.17582	.07692	.03571	.01753	.00905	.00488
	3		1.	.67033	.40659	.24176	.14480	.08824	.05489
	4			1.	.84615	.65385	.48416	.35219	.25494
	5				1.	.94231	.84050	.72172	.60503
	6					1.	.98303	.94344	.88509
	7						1.	.99623	.98512
	8							1.	.99940
	9								1.
	10								
	11								
12	1	.14286	.03226	.00862	.00275	.00096	.00041	.00016	.00007
	2	1.	.38710	.15517	.06314	.02788	.01377	.00626	.00317
	3		1.	.62931	.36538	.20481	.12353	.06657	.03938
	4			1.	.81868	.60288	.37667	.29339	.20210
	5				1.	.92019	.77764	.65523	.52754
	6					1.	.97327	.92058	.84524
	7						1.	.99326	.97542
	8							1.	.99866
	9								1.
	10								
	11								
	12								
13	1	.14286	.02857	.00714	.00210	.00070	.00026	.00010	.00004
	2	1.	.37143	.13571	.05252	.02171	.00955	.00444	.00217
	3		1.	.60714	.32983	.17577	.09469	.05212	.02726
	4			1.	.79202	.56092	.37848	.25077	.16379
	5				1.	.09756	.76161	.60836	.47097
	6					1.	.96594	.89443	.79863
	7						1.	.98978	.96246
	8							1.	.99756
	9								1.
	10								
	11								
	12								
	13								

Table of P{2r ≤ 2k} (continued)

N_2	k	$N_1=10$	$N_1=11$	$N_1=12$	$N_1=13$	$N_1=14$	$N_1=15$
	1	.00011					
	2	.00454					
	3	.05123					
	4	.24233					
10	5	.58560					
	6	.87224					
	7	.98119					
	8	.99892					
	9	.99989					
	10	1.					
	1	.00006	.00003				
	2	.00274	.00159				
	3	.03489	.02264				
	4	.18492	.13491				
11	5	.50000	.40997				
	6	.81508	.74004				
	7	.96511	.93651				
	8	.99726	.99264				
	9	.99994	.99966				
	10	1.	.99997				
	11		1.				
	1	.00003	.00002	.00001			
	2	.00171	.00094	.00053			
	3	.02382	.01477	.00923			
	4	.14006	.09772	.06795			
	5	.41863	.33001	.25568			
12	6	.75864	.66999	.59091			
	7	.94436	.90228	.85373			
	8	.99427	.98523	.97118			
	9	.99980	.99906	.99727			
	10	1.	.99998	.99990			
	11		1.	.99999			
	12			1.			
	1	.00002	.00001	.00000	.00000		
	2	.00111	.00059	.00032	.00018		
	3	.01703	.01011	.00614	.00381		
	4	.10991	.07356	.04977	.03406		
	5	.36068	.27346	.20682	.15657		
	6	.69505	.59328	.50000	.41791		
13	7	.91796	.85981	.79318	.72281		
	8	.98961	.97403	.95023	.91882		
	9	.99956	.99783	.99386	.98688		
	10	1.	.99994	.99968	.99898		
	11		1.	1.	1.		
	12			1.	1.		
	13				1.		

Table of P{2r ≤ 2k} (continued)

N_2	k	$N_1=2$	$N_1=3$	$N_1=4$	$N_1=5$	$N_1=6$	$N_1=7$	$N_1=8$	$N_1=9$
	1	.12500	.25000	.00581	.00163	.00051	.00018	.00006	0.00003
	2	1.00000	.35000	.12209	.04412	.01749	.00722	.00299	.00149
	3		1.00000	.57558	.29902	.15123	.07765	.03768	.02207
	4			1.	.76634	.52058	.33586	.19696	.13483
	5				1.	.88837	.72319	.51496	.41688
	6					1.	.95558	.87718	.75421
14	7						1.	.98621	.94762
	8							1.	.99597
	9								1.
	10								
	11								
	12								
	13								
	14								
	1	.12500	.02174	.00490	.00129	.00039	.00013	.00005	.00002
	2	1.	.32609	.10784	.03737	.01392	.00555	.00238	.00105
	3		1.	.55392	.27191	.13148	.06426	.03273	.01665
	4			1.	.74098	.48337	.29910	.18450	11018
	5				1.	.87046	.68632	.51837	.36739
	6					1.	.94465	.83647	.71039
	7						1.	.97956	.93086
15	8							1.	.99385
	9								1.
	10								
	11								
	12								
	13								
	14								
	15								
	1	.11111	.01961	.00408	.00103	.00029	.00009	.00003	.00001
	2	1.	.31373	.09796	.03199	.01147	.00424	.00176	.00075
	3		1.	.52653	.24871	.11445	.05258	.02573	.01274
	4			1.	.71827	.45013	.28050	.15574	.09069
	5				1.	.85173	.65761	.46731	.32454
	6					1.	.93416	.81038	.66752
	7						1.	.97359	.91251
16	8							1.	.99125
	9								1.
	10								
	11								
	12								
	13								
	14								
	15								
	16								

N_2	k	$N_1=10$	$N_1=11$	$N_1=12$	$N_1=13$	$N_1=14$	$N_1=15$	$N_1=16$
14	1	.00001	.00000	.00000	.00000	.00000		
	2	.00073	.00037	.00020	.00011	.00006		
	3	.01218	.00693	.00404	.00242	.00148		
	4	.08572	.05505	.03574	.02359	.01567		
	5	.30612	.22348	.16249	.11887	.08661		
	6	.63690	.52665	.42873	.34755	.27817		
	7	.88879	.81539	.73292	.65245	.57002		
	8	.98336	.95975	.92526	.88113	.83131		
	9	.99911	.99584	.98863	.97641	.95899		
	10	1.	.99985	.99921	.99758	.99448		
	11		1.	.99998	.99989	.99964		
	12			1.	1.00000	.99999		
	13				1.	1.		
	14					1.		
15	1	.00001	.00000	.00000	.00000	.00000	.00000	
	2	.00049	.00024	.00012	.00006	.00003	.00002	
	3	.00889	.00483	.00272	.00156	.00092	.00055	
	4	.06765	.04158	.02605	.01653	.01065	.00696	
	5	.26156	.18306	.12873	.09064	.06417	.04572	
	6	.57964	.46602	.36857	.28826	.22674	.17491	
	7	.85665	.76919	.67661	.58470	.50000	.42407	
	8	.97537	.94243	.89664	.83878	.77526	.70882	
	9	.99845	.99296	.98221	.96229	.93583	.90261	
	10	1.	.99969	.99840	.99523	.98935	.98012	
	11		1.	.99995	.99972	.99908	.99774	
	12			1.	.99999	.99997	.99987	
	13				1.	1.	1.	
	14					1.	1.	
	15						1.	
16	1	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	2	.00034	.00016	.00008	.00004	.00002	.00001	.00000
	3	.00650	.00342	.00185	.00103	.00058	.00034	.00020
	4	.05330	.03169	.01913	.01172	.00729	.00461	.00295
	5	.22163	.15023	.10204	.06947	.04750	.03277	.02277
	6	.53034	.41148	.31487	.23886	.18022	.13598	.10269
	7	.82425	.72249	.61888	.52117	.43300	.35717	.29296
	8	.96604	.92242	.86321	.79341	.71867	.64283	.56942
	9	.99753	.98907	.97178	.94465	.90828	.86402	.81406
	10	1.	.99943	.99713	.99170	.98199	.96723	.94726
	11		1.	.99989	.99940	.99807	.99539	.99085
	12			1.	1.	.99990	.99966	.99911
	13				1.	1.	1.	.99996
	14					1.	1.	1.
	15						1.	1.
	16							1.

N_2	k	$N_1=2$	$N_1=3$	$N_1=4$	$N_1=5$	$N_1=6$	$N_1=7$	$N_1=8$	$N_1=9$
17	1	.11111	.01754	.00351	.00084	.00023	.00007	.00002	.00001
	2	1.	.29824	.08772	.02757	.00934	.00340	.00132	.00046
	3		1.	.50877	.22807	.10048	.04500	.02073	.00843
	4			1.	.69591	.41946	.23917	.13399	.06421
	5				1.	.83413	.61778	.42847	.39051
	6					1.	.92067	.78185	.68055
	7						1.	.96695	.90844
	8							1.	.98983
	9								1.
	10								
	11								
	12								
	13								
	14								
	15								
	16								
	17								
18	1	.10000	.01563	.00299	.00068	.00018	.00005	.00002	.00001
	2	1.	.28125	.08060	.02392	.00783	.00270	.00101	.00040
	3		1.	.48657	.20984	.08861	.03807	.01685	.00771
	4			1.	.67464	.39181	.21487	.11593	.06255
	5				1.	.81530	.58617	.39306	.25450
	6					1.	.90806	.75357	.58722
	7						1.	.95944	.87239
	8							1.	.98442
	9								1.
	10								
	11								
	12								
	13								
	14								
	15								
	16								
	17								
	18								

N_2	k	N_1=10	N_1=11	N_1=12	N_1=13	N_1=14	N_1=15	N_1=16	N_1=17	N_1=18
17	1	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	
	2	.00023	.00011	.00005	.00002	.00001	.00001	.00000	.00000	
	3	.00484	.00245	.00128	.00068	.00038	.00021	.00012	.00007	
	4	.04248	.02436	.01419	.00832	.00506	.00309	.00192	.00121	
	5	.18927	.12403	.08131	.05300	.03548	.02371	.01598	.01087	
	6	.48284	.36323	.26926	.19600	.14501	.10615	.07781	.05720	
	7	.79039	.67648	.56461	.45817	.37450	.30047	.23974	.19067	
	8	.95515	.90023	.82831	.73906	.66136	.57808	.50000	.42902	
	9	.99634	.98413	.96016	.91461	.87650	.82098	.76026	.69716	
	10	1.	.99905	.99532	.97703	.97212	.95053	.92219	.88784	
	11		1.	.99980	.99888	.99646	.99175	.98402	.97278	
	12			1.	.99997	.99978	.99925	.99808	.99595	
	13				1.	1.	.99997	.99988	.99966	
	14					1.	1.	1.	.99999	
	15						1.	1.	1.	
	16							1.	1.	
	17								1.	
18	1	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	2	.00017	.00008	.00003	.00002	.00001	.00000	.00000	.00000	.00000
	3	.00365	.00190	.00090	.00047	.00025	.00014	.00007	.00004	.00003
	4	.03413	.02009	.01063	.00609	.00355	.00210	.00126	.00078	.00048
	5	.16208	.10926	.06511	.04150	.02665	.01726	.01122	.00746	.00498
	6	.43936	.34110	.23038	.16427	.11675	.08296	.05871	.04226	.03030
	7	.75618	.67231	.51367	.40980	.32268	.25189	.19437	.15146	.11712
	8	.94294	.86831	.79193	.69919	.60584	.51735	.44039	.36595	.30445
	9	.99480	.97674	.94651	.90015	.84180	.77544	.70687	.63407	.56500
	10	1.	.99842	.99289	.98053	.95978	.93029	.89341	.84856	.79950
	11		1.	.99964	.99807	.99411	.98660	.97481	.95776	.93594
	12			1.	.99993	.99957	.99854	.99640	.99256	.98658
	13				1.	.99999	.99992	.99972	.99924	.99827
	14					1.	1.	.99999	.99996	.99987
	15						1.	1.	1.	1.
	16							1.	1.	1.
	17								1.	1.
	18									1.

Table of P{2r ≤ 2k} (continued).

N_2	k	$N_1=2$	$N_1=3$	$N_1=4$	$N_1=5$	$N_1=6$	$N_1=7$	$N_1=8$	$N_1=9$
	1	.10000	.01429	.00260	.00056	.00014	.00004	.00001	.00000
	2	1.	.27143	.07273	.02089	.00649	.00217	.00078	.00030
	3		1.	.47013	.19368	.07849	.03241	.01384	.00608
	4			1.	.65443	.36646	.19368	.09840	.05240
	5				1.	.79842	.55652	.35955	.22607
	6					1.	.89518	.72517	.55025
	7						1.	.95150	.85128
	8							1.	.98029
19	9								1.
	10								
	11								
	12								
	13								
	14								
	15								
	16								
	17								
	18								
	19								
	1	.09091	.01299	.00224	.00047	.00011	.00003	.00001	.00000
	2	1.	.25974	.06726	.01834	.00524	.00176	.00061	.00022
	3		1.	.45067	.17921	.06617	.02776	.01139	.00484
	4			1.	.63500	.37781	.17507	.08781	.04413
	5				1.	.79217	.52861	.33216	.20126
	6					1.	.88215	.69883	.51552
	7						1.	.94317	.82978
	8							1.	.97568
	9								1.
20	10								
	11								
	12								
	13								
	14								
	15								
	16								
	17								
	18								
	19								
	20								

N_2	k	$N_1=10$	$N_1=11$	$N_1=12$	$N_1=13$	$N_1=14$	$N_1=15$	$N_1=16$	$N_1=17$
19	1	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	2	.00012	.00005	.00002	.00001	.00000	.00000	.00000	.00000
	3	.00278	.00131	.00064	.00032	.00017	.00009	.00005	.00003
	4	.02759	.01475	.00800	.00445	.00252	.00145	.00085	.00050
	5	.13923	.08533	.05221	.03236	.02015	.01267	.00805	.00517
	6	.39974	.28296	.19663	.13655	.09422	.06503	.04502	.03129
	7	.72227	.58881	.46484	.36228	.27763	.21088	.15944	.12021
	8	.92961	.85096	.75569	.65251	.55274	.46092	.38012	.31077
	9	.99296	.97111	.93130	.87422	.80493	.72832	.64984	.57278
	10	1.	.99782	.98984	.97275	.94504	.90685	.85963	.80568
	11		1.	.99942	.99694	.99089	.97978	.96261	.93907
	12			1.	.99987	.99923	.99746	.99382	.98758
	13				1.	.99998	.99984	.99942	.99846
	14					1.	1.	.99997	.99989
	15						1.	1.	1.
	16							1.	1.
	17								1.
	18								
	19								
20	1	.00000	.00000	.00000	.00000	.00000	.00000	.00000	.00000
	2	.00010	.00004	.00002	.00001	.00000	.00000	.00000	.00000
	3	.00237	.00097	.00046	.00022	.00011	.00006	.00003	.00002
	4	.02495	.01162	.00613	.00327	.00181	.00102	.00057	.00033
	5	.13329	.07121	.04239	.03251	.01541	.00947	.00578	.00361
	6	.29325	.25002	.16928	.12023	.07611	.05174	.03445	.02331
	7	.65437	.54802	.42305	.32493	.23925	.16764	.13001	.09551
	8	.90591	.82474	.71763	.61004	.50435	.40312	.32968	.26313
	9	.98974	.96310	.91400	.84763	.76945	.67784	.59590	.51456
	10	1.	.99693	.98601	.96378	.93146	.87930	.82366	.76040
	11		1.	.99911	.99546	.99037	.97088	.94790	.91684
	12			1.	.99978	.99872	.99585	.99025	.98084
	13				1.	.99995	.99970	.99894	.99725
	14					1.	.99999	.99994	.99977
	15						1.	1.	.99999
	16							1.	1.
	17								1.
	18								
	19								
	20								

Table of P{2r $\leq$ 2k} (continued).

N_2	k	N_1=18	N_1=19	N_1= 20
19	1	.00000	.00000	
	2	.00000	.00000	
	3	.00001	.00001	
	4	.00031	.00020	
	5	.00335	.00236	
	6	.02186	.01648	
	7	.09057	.07330	
	8	.25254	.14595	
	9	.50000	.39148	
	10	.74746	.66429	
	11	.90943	.86518	
	12	.97814	.96258	
	13	.99665	.99318	
	14	.99969	.99923	
	15	.99999	.99995	
	16	1.	1.	
	17	1.	1.	
	18	1.	1.	
	19		1.	
20	1	.00000	.00000	.00000
	2	.00000	.00000	.00000
	3	.00001	.00001	.00000
	4	.00019	.00012	.00007
	5	.00227	.00146	.00095
	6	.01575	.01085	.00319
	7	.06968	.05157	.03383
	8	.20740	.16499	.12633
	9	.43693	.37293	.31131
	10	.69596	.62707	.56001
	11	.87959	.83501	.78610
	12	.96722	.94843	.92484
	13	.99418	.98915	.98176
	14	.99937	.99854	.99708
	15	.99996	.99988	.99971
	16	1.	1.	.99998
	17	1.	1.	1.
	18	1.	1.	1.
	19		1.	1.
	20			1.

4. Numerical example.

To illustrate the use of the numerical table in testing for randomness of an arrangement under the null hypothesis that two sets of observations are from the same distribution, let us consider the following example. Watson (1962) studied a problem related to the migration of birds. In his data the measurements are given only to the nearest 5°, as follows:

Control group ($N_1 = 12$): 50, 290, 300, 300, 305, 320, 330, 330, 335, 340, 340, 355

Experimental group ($N_2 = 14$): 70, 155, 190, 195, 215, 235, 235, 240, 255, 260, 290, 300, 300, 300.

In this example, unfortunately, due to grouping some ties occur between values from the two different sets of observations. However, breaking up the ties we can give upper and lower bounds for the number of ties in favor or against the null hypothesis.

Following Watson, let us first change 290, 300, 300 of the control group into 285, 295, 295. Then the number of runs observed is 6. From the Table we find $P(2r \leq 6) = 0.0040$ such that the two samples are significantly different at an often-used level of 0.01.

Second, it is easy to see that for various ways in breaking up the ties the lower bound is 4 runs, the upper bound 8 runs. Thus we obtain

$$\text{Prob}\,\{2r \leq 4\} = 0.0002 \leq \text{Prob}\{2r \leq \text{ the actual number of runs}\}$$
$$\leq \text{Prob}\{2r \leq 8\} = 0.0357\ .$$

From this we conclude that at the 5% level the two groups differ significantly.

I am indebted to Professor E. Batschelet for suggesting to work in the field of circular distribution, and to Mr. J. F. Gilroy, S.J., for his help in preparing this manuscript, (The Catholic University of America). I wish also to thank Professor G. S. Watson (Johns Hopkins University) for referring me to Barton and David's work (1958, 1962).

References

Barton, D. E. and David, F. N., (1958). Runs in a ring. Biometrika, 45, 572-578.

Barton, D. E. and David, F. N., (1962). Combinatorial chance. Hafner Publishing Co.

Curray, J. R. (1956). The analysis of two-dimensional orientation data. Jour. Geology 64, 117-131.

Durand, D. and Greenwood, J. A. (1957). Random unit vectors. II. Usefulness of Gram-Charlier and related series in approximating distributions. Ann. Math. Stat. 28, 978-986.

Durand, D. and Greenwood, J. A. (1958). Modifications of the Rayleigh test for uniformity in analysis of two-dimensional orientation data. Jour. Geology. 66, 229-238.

Greenwood, J. A. and Durand, D. (1955). The distribution of the length and components of the sum of n random unit vectors. Ann. Math. Stat. 26, 233-246.

Gumbel, E. J. Greenwood, J. A. and Durand D. (1953). The circular normal distribution: Theory and tables. Jour. Ann. Stat. Assoc. 48, 131-152.

Kuiper, N. H. (1960). Tests concerning random points on a circle. Indag. Math. 22, No. 1, 38-47.

Kuiper, N. H., (1960). Random variables and random vectors. Bull. Inst. Agronomique et Station de Recher. de Gemblonx. 1, 344-355.

MacMahon, M. P. A. (1916). Combinatory analysis, v. 1 & 2, Cambridge University Press.

Mood, A. M. (1940). The distribution theory of "Runs". Ann. Math. Stat. 11, 367-392.

Riordan, J., (1958). An introduction to combinatorial analysis. John Wiley & Sons.

Schmidt-Koenig, K. (1960). Internal clocks and homings. Cold Spring Harbor symposium on Quantitative Biology, 25, 389-393.

Stephens, M. A. (1962). Exact and approximate tests for directions (I). Biometrika, 49, 463-477.

Stevens, W. L. (1939). Distribution of groups in a sequence of alternatives. Ann. Eugenics, 9, 10-17.

Swed, F. S. and Eisenhart, C. (1943). Tables for testing randomness of grouping in a sequence of alternatives. Ann. Math. Stat. 14, 66-87.

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are from the same population. Ann. Math. Stat. 11, 147-162.

Watson, G. S. and Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere. Biometrika 43, 344-352.

Watson, G. S. (1961). Goodness-of-fit on a circle. Biometrika 48, 109-114.

Watson, G. S. (1962). Goodness-of-fit tests on a circle (II). Biometrika 49, 57-63.